Published in final edited form as:

Psychophysiology. 2019 December ; 56(12): e13454. doi:10.1111/psyp.13454.

Physiological Indices of Challenge and Threat: A Data-Driven Investigation of Autonomic Nervous System Reactivity During an Active Coping Stressor Task

Jolie B. Wormwood[1,*], **Zulqarnain Khan**[2,*], **Erika Siegel**[3], **Spencer K. Lynn**[2,4], **Jennifer Dy**[2], **Lisa Feldman Barrett**[2,5], **Karen S. Quigley**[2,6]

[1]University of New Hampshire, Durham, NH

[2]Northeastern University, Boston, MA

[3]University of California San Francisco, San Francisco, CA

[4]Charles River Analytics, Inc., Cambridge, MA

[5]Massachusetts General Hospital, Cambridge, MA

[6]Edith Nourse Rogers Memorial (VA) Medical Center, Bedford, MA

Abstract

We utilized a data-driven, unsupervised machine learning approach to examine patterns of peripheral physiological responses during a motivated performance context across two large, independent data sets, each with multiple peripheral physiological measures. Results revealed that patterns of cardiovascular response commonly associated with 'challenge' and 'threat' states emerged as two of the predominant patterns of peripheral physiological responding within both samples, with these two patterns best differentiated by reactivity in cardiac output (CO), pre-ejection period (PEP), interbeat interval (IBI), and total peripheral resistance (TPR). However, we also identified a third, relatively large group of apparent physiological "non-responders" who exhibited minimal reactivity across all physiological measures in the motivated performance context. This group of "non-responders" was best differentiated from the others by minimal increases in electrodermal activity. We discuss implications for identifying and characterizing this third group of individuals in future research on physiological patterns of challenge and threat.

Keywords

Threat; Challenge; Biopsychosocial Model; Motivated Performance; Machine Learning; Impedance Cardiography

Correspondence: Correspondence concerning this article should be addressed to Jolie Wormwood, Department of Psychology, 422 McConnell Hall, University of New Hampshire, 15 Academic Way, Durham, NH 03864; jolie.wormwood@unh.edu; or to Zulqarnain Khan, Department of Electrical and Computer Engineering, 409 Dana Research Center, Northeastern University, 360 Huntington Avenue, Boston, MA 02115; khan.zu@husky.neu.edu.

*Authors contributed equally

1 Introduction

Motivated performance contexts are those situations that are goal-relevant to the performer, require an instrumental response, and require active engagement (i.e., active coping) rather than passive intake (i.e., passive coping; Blascovich & Mendes, 2000). Many commonly-used cognitive stressor tasks (e.g., problem-solving tasks, Stroop tasks) and socio-evaluative stressor tasks (e.g., public speaking, interpersonal negotiation) are motivated performance contexts when they meet these three criteria. Social psychologists interested in motivated performance contexts have differentiated two profiles of biopsychological responding termed “challenge” and “threat” (for discussions, see Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996; Seery, 2011; Tomaka, Blascovich, Kelsey, & Leitten, 1993). Individuals experiencing challenge evaluate their resources as sufficient for coping with the demands of a stressful situation or task. Challenge is associated with a pattern of cardiovascular reactivity thought to be beneficial for performance wherein the heart pumps more blood per unit of time and circulates it to the periphery more efficiently. This is typically indexed by increases in cardiac output (CO, volume of blood circulated per minute), decreases in pre-ejection period (PEP, the time between the electrical stimulus initiating ventricular contraction and opening of the aortic valve), decreases in total peripheral resistance (TPR, the extent of overall constriction in the peripheral vasculature), and decreases in interbeat interval (IBI, the time between consecutive heartbeats). In contrast, individuals experiencing threat evaluate their personal resources as insufficient for coping with the demands of a situation or task. Threat is associated with more modest changes from baseline in cardiac reactivity compared to challenge, and is accompanied by minimal change or increases in vasoconstriction in the periphery. This pattern of cardiovascular activity is thought to be detrimental to performance because there is reduced blood flow to the periphery to support action. Threat is typically indexed by small increases in CO, and small decreases in PEP and IBI along with modest increases (or no change) in TPR (see, for example, Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996; Dienstbier, 1989; Mendes, Blascovich, Hunter, Lickel, & Jost, 2007; Quigley, Barrett, & Weinstein, 2002; Tomaka et al., 1993; Tomaka, Blascovich, Kibler, & Ernst, 1997).

This theoretical framework has proven useful in describing and predicting behavior across a wide variety of motivated performance tasks (e.g., Brimmel, Parker, Furley, & Moore, 2018; Zilka, Rahimi, & Cohen, 2019; Streamer, Seery, Kondrak, Lamarche, & Saltsman, 2017; Jamieson, Mendes, Blackstock, & Schmader, 2010; Mendes, Blascovich, Major, & Seery, 2001; Alter, Aronson, Darley, Rodriguez, & Ruble, 2010; Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001; Drach-Zahavy & Erez, 2002; Harvey, Nathens, Bandiera, & LeBlanc, 2010; Mendes, Blascovich, Lickel, & Hunter, 2002; Mendes, Gray, Mendoza-Denton, Major, & Epel, 2007; Quigley et al., 2002; Skinner & Brewer, 2002). However, early adoption of a dichotomous challenge/threat approach led some investigators to assume that the physiological reactivity of all individuals in a sample can be simply categorized into one of these two orthogonal biopsychological states. Indeed, in the vast majority of studies utilizing this approach, analyses begin by designating two groups of individuals for comparison, based either on the experimental design (e.g., random assignment to one of two experimental conditions) or on a median split (or other dichotomization) of the self-reported appraisals of perceived task demands and coping resources. Researchers then compare physiological reactivity across the two groups, where 'reactivity' is defined using the mean of a physiological variable during a task after subtracting the mean during a baseline epoch. Using these methods, there is no way to detect potential patterns of biopsychological reactivity that are not challenge or threat, since other patterns are by default subsumed into one of these two categories.

In the early development of the challenge/threat framework, challenge and threat states were linked most closely to cardiovascular function (Dienstbier, 1989; Tomaka et al., 1993; 1997), and investigators quickly narrowed their experimental focus almost exclusively to cardiovascular variables (e.g., CO, PEP, IBI, TPR). Thus, more recently there has been little use of non-cardiovascular measures that might be related to patterns of physiological reactivity during active coping other than challenge and threat. To address this issue, and in a reversal of the usual approach to studying associations between cardiovascular reactivity and self-reported appraisals, we measured a broad set of different peripheral physiological measures during a commonly used motivated performance task and then used a data-driven analysis to discover, rather than stipulate, the number of distinct patterns of peripheral physiological reactivity across individuals. Then, we explored whether any discovered physiological patterns were associated with theoretically-expected levels of self-reported stress and coping.

In the present investigation, we utilized an unsupervised machine learning approach to investigate peripheral physiological reactivity in a motivated performance context (i.e., during mental math tasks). This approach allowed us to identify and characterize the predominant patterns of physiological reactivity within a motivated performance context in a completely data-driven way, and then to quantify the extent to which each peripheral physiological variable contributed to these patterns of reactivity. In addition, we employed this unsupervised machine learning approach across two large, independent samples, allowing us to assess the consistency and reliability of our findings. These independent data sets each included a large number of peripheral physiological measures, including indices unrelated to cardiovascular reactivity (e.g., facial muscle movement, respiration, and electrodermal activity) which typically have not been included in research on challenge and threat states. We predicted that two of the predominant patterns of physiological reactivity identified across samples would be consistent with the patterns of cardiovascular reactivity commonly associated with challenge and threat states, replicating existing literature and providing strong empirical support for this theoretical framework. Here, for the first time using such a broad sampling of physiological variables, we explored whether additional patterns of peripheral physiological reactivity beyond those associated with challenge and threat were apparent within a motivated performance context and, if so, which specific physiological features best differentiated any new patterns of physiological reactivity from those of challenge and threat.

2 Method

2.1 Overview

Two independently collected datasets from our lab were used in the analyses and designated as Dataset 1 and Dataset 2.

2.2 Participants

Dataset 1 was drawn from a sample of 260 participants (100 males, 159 females, 1 not reported) between the ages of 18 and 65 ($M = 24.8$, $SD = 10.2$) with BMIs ranging from 13.0 to 48.5 ($M = 24.9$, $SD = 5.4$). Dataset 2 was drawn from a sample of 300 participants (115 males, 171 females, 14 not reported) between the ages of 18 and 55 ($M = 23.4$, $SD = 8.0$) with BMIs ranging from 15.2 to 44.6 ($M = 24.4$, $SD = 4.7$). Participants for both datasets were recruited from the greater Boston area through fliers on college campuses, and via advertisements on craigslist.com and in a local newspaper. Participants were excluded for skin allergies, sensitive skin, a history of cardiovascular illness or stroke, asthma, other chronic medical conditions, or if they self-reported any current mental illness. Participants were also excluded if they were currently taking medications to treat ADHD, insomnia, anxiety, high blood pressure, rheumatoid arthritis, epilepsy/seizures, cold/flu, or hay fever/allergies (not including nasal sprays or other non-autonomically active medications). Eligible participants in Dataset 1 were asked to refrain from caffeine (96% compliance), tobacco (96% compliance), diet pills or sleeping pills (99% compliance), and alcohol (99% compliance) for 24 hours prior to the experiment. Eligible participants in Dataset 2 were asked to refrain from caffeine and tobacco for 12 hours prior to the experiment (100% compliance). Participants in Dataset 1 were required to be native English speakers. All participants completed a mental mathematics (math) task embedded within a larger set of experimental tasks (for details, see supplemental online materials). Participants in both datasets received $5 per half hour of participation, although some participants received additional compensation because they were completing tasks as part of a multiple-visit experiment (for details, see supplemental online materials).

Only participants with complete peripheral physiological data were included in the clustering analyses (i.e., participants needed to have artifact-free, valid data (i.e., visually inspected and deemed to be good quality) for all 12 (Dataset 1) or 10 (Dataset 2) peripheral physiological features for the baseline and for all minutes of a mental math task). Thus, the analyses for the present investigation include only 165 participants from Dataset 1 (58 males, 106 females, 1 not reported; $M_{age} = 23.57$, $SD_{age} = 9.12$; $M_{BMI} = 24.08$, $SD_{BMI} = 4.49$), and only 130 participants from Dataset 2 (50 males, 80 females; $M_{age} = 22.91$, $SD_{age} = 7.64$; $M_{BMI} = 23.56$, $SD_{BMI} = 4.78$). Table 1 presents demographic, anthropometric, and baseline physiological activity for both the included and excluded participants in both datasets.

2.3 Procedure

All participants completed a mental math task embedded within a larger experimental session. Details concerning aspects of the experiments beyond the mental math task, including information on the procedure and tasks unrelated to the current investigation, can be found in the supplemental online materials. For both datasets, participants were instrumented for physiological recordings and completed a 2-minute (Dataset 1) or 5-minute (Dataset 2) baseline recording period prior to completing the mental math task. To ensure this resting baseline was not confounded with anticipatory stress related to the upcoming mental math task, the math task was not described in detail prior to the baseline.

2.3.1 Mental Mathematics Task.—Aspects of the procedure for the mental math task differed across Dataset 1 and Dataset 2. Detailed procedures for both can be found in the supplemental online materials, but we outline the major shared components here. For both datasets, the participant was informed that s/he would complete a mental math task in which s/he would perform serial subtractions aloud in front of an experimenter, who would be recording and evaluating his or her responses. The participant was asked to work as quickly and accurately as possible, and to refrain from commenting on the task or his or her performance until the task was complete. Experimenters were trained not to provide any positive feedback (e.g., no smiling or nodding) during the task. In Dataset 1, participants completed four minutes of mental math across three different sets of 3 or 4 digit seed numbers from which they subtracted a smaller subtrahend. Participants completed serial subtractions for 1 minute for the first seed number, for 1 minute for the second seed number, and for 2 minutes for the third and final seed number. No feedback was provided during the first two minutes of mental math, but during the final two minutes of mental math in Dataset 1, participants were informed when they were incorrect and were told the last seed number from which they had made a correct subtraction. In Dataset 2, participants completed five minutes of mental math, completing serial subtractions for 1 minute for each of five different seed numbers. For each of the five minutes, participants were informed when they were incorrect, told the last seed number from which they had made a correct subtraction, reminded of the subtrahend, and asked to continue. For both datasets, the seed numbers were adjusted across the minutes of the task based on participants' performance, with the goal of keeping the serial subtractions moderately stressful for all participants by adjusting the task difficulty either based on their first minute's performance (Dataset 1) or their performance during the prior minute of the task (Dataset 2). In addition, participants in both datasets made appraisal ratings (e.g., how stressful the upcoming task would be and how well they thought they would cope) immediately before the first minute of mental math and immediately following the final minute of mental math. The specific ratings differed across datasets (see below and in the supplemental online materials), but in both datasets the appraisals measured the perceived task demands (e.g., ratings of how stressful they expected or perceived the task to be) and the perceived resources to meet those demands (e.g., ratings of how well they expected to or did cope with the task). For Dataset 1 only, participants also made a third set of appraisal ratings in the middle of the mental math task, immediately before the third minute of mental math began.

2.4 Physiological Measures

2.4.1 Physiological Data Acquisition.—For Dataset 1, we recorded the electrocardiogram (ECG), the impedance cardiogram (IC), continuous blood pressure (BP), respiration, electrodermal activity (EDA), and measures of facial muscle activity over the corrugator supercilii (CORR) and zygomaticus major (ZYGO) muscle regions using facial electromyography (fEMG). All physiological measures in Dataset 1 were sampled at 1000 Hz using BioLab v. 3.0.13 (Mindware Technologies; Gahanna, OH), and were acquired on a BioNex 8-Slot Chassis (Model 50–3711-08). For Dataset 2, we recorded ECG, IC, BP and EDA, with BP collected intermittently instead of continuously. All physiological measures in Dataset 2 were sampled at 500 Hz using a mobile impedance cardiograph (Model 50–2303-00; Mindware Technologies, Gahanna, OH). Respiration was measured via a piezoelectric belt placed around the lower-chest/upper-abdomen (Mindware Technologies; Model 50–4504-00; Dataset 1) or was derived from the IC signal (Dataset 2). ECG was obtained using pre-gelled Ag/AgCl sensors in a modified lead II configuration. fEMG measures (in Dataset 1) were obtained via reusable 4mm Ag/AgCl electrodes (Mindware Technologies; Gahanna, OH) filled with a high-conductivity gel over the zygomaticus major and corrugator supercilii muscle regions on the right side of the participant's face. A reference electrode was placed in the middle of the forehead. Before sensor attachment, each site was cleaned with alcohol and exfoliated using a semi-abrasive lotion with the goal of achieving a contact impedance <10 KOhms (M = 9.57, SD = 13.07 for final sample included in analyses). The IC was acquired using a four-spot electrode configuration (see Qu, Zhang, Webster, & Tompkins, 1986) using pre-gelled disposable Ag/AgCl electrodes. The inner (recording) electrodes were placed on the participant's chest: one at the base of the neck at the top of the sternum and the other at the level of the xiphisternal junction. The outer (source) electrodes were placed on the participant's back along the midline approximately 4 cm above and below the inner recording electrodes (respectively, roughly over the fourth cervical vertebra and the ninth thoracic vertebra). The source electrodes passed a 4 mA, 100 kHz alternating current across the thorax. Electrodermal activity (EDA) was recorded from the palmar surface (right hand for Dataset 1, dominant hand for Dataset 2) with sensors on the thenar and hypothenar eminences using disposable, Ag/AgCl (11mm diameter; isotonic paste) electrodes (Biopac Systems, Inc.; Goleta, CA). When needed, a small amount of additional isotonic paste was added to the electrode surface. Prior to electrode placement, participants washed their hands with warm water. Blood pressure in Dataset 1 was recorded continuously via a Continuous Noninvasive Arterial Pressure monitor (CNAP Monitor 500AT; CNSystems; Medizintechnik, AG, Austria). Continuous recordings were obtained from finger cuffs placed on participants' left middle and index fingers, and these continuous readings were calibrated against intermittent non-invasive blood pressure measurements from a cuff placed around the participant's right arm. Blood pressure in Dataset 2 was recorded intermittently via an arm cuff placed around the arm of the participant's non-dominant hand. Three blood pressure readings were taken during the baseline period (and averaged), and three blood pressure readings were taken during the mental math task: before the end of the first minute, during the second minute, and before the end of the fifth minute of mental math. In our analyses with Dataset 2, we used the first BP reading as our measure for the first two minutes of the mental math task, the second reading for our measure of the third and fourth minutes of the mental math task, and the third reading for the final minute of the mental math task.

2.4.2 Physiological Signal Processing and Feature Extraction.—All

physiological data were processed in software programs for scoring physiological signals from Mindware Technologies LTD (v3.0.25). Dependent variables (hereafter features) from each signal were calculated as averages for each one-minute segment of the baseline and trials of the mental math task; data from each baseline minute were averaged to obtain a single baseline estimate for each feature. Using all available physiological data in each dataset, we extracted 12 (Dataset 1) or 10 (Dataset 2) peripheral physiological features in total for each segment, including: respiration rate (RR); interbeat interval (IBI: average R-R interval); number of skin conductance responses (SCRs: defined as a .01 μS increase in skin conductance); skin conductance level (SCL: average skin conductance, excluding SCRs from onset to half-recovery time); mean arterial pressure (MAP); total peripheral resistance (TPR); and, for Dataset 1 only, root mean square (or quadratic mean) for each fEMG channel (i.e., ZYGO and CORR). We also extracted respiratory sinus arrhythmia (RSA: calculated using spectral analysis), which reflects high-frequency variability in IBI (i.e., in the .12-.40 Hz range) and provides an estimate of parasympathetic control of the heart. From the impedance signal, we derived one volumetric measure of cardiac function: cardiac output (CO) in L/min, and two systolic time intervals: pre-ejection period (PEP; an estimate of sympathetic cardiac control) and left ventricular ejection time (LVET). Stroke volume was not derived as a feature to be included in clustering analyses because of its interdependence with other features (i.e., heart rate, CO, TPR). All raw data were subjected to manual, visual inspection by trained research assistants, and segments with movement artifact or data of insufficient quality for valid feature extraction were removed from all analyses. Finally, average baseline values for each feature were subtracted from the values for each feature during each minute of the mental math task. Thus, the analyses utilize difference scores for all physiological features. More detailed information on the data processing for specific physiological features is available in the supplemental online materials.

3 Analysis

The goal of the present investigation was to define similar groups of participants based on their physiological data without imposing any a priori assumptions about the variables that would distinguish those groups. Then, we examined whether interpretable patterns were found as a result of that grouping. This is a clustering problem which can be answered using a data-driven approach known as unsupervised learning (i.e., discovering sub-groups within samples based solely on the data without any pre-determined labels). Unsupervised learning contrasts with supervised learning approaches in which models are trained based on already labeled data.

In machine learning, the objective of clustering is to find groups such that members within a group are ‘similar’ to each other, while being ‘different’ from members of other groups (i.e., minimize intra-cluster variance, and maximize inter-cluster variance). Various clustering algorithms have been developed, which differ in terms of complexity, types of priors that can be incorporated, types of groups that the algorithm discovers, and the intermediate outputs that can be obtained (Barbakh, Wu, & Fyfe, 2009).

3.1 Hierarchical Clustering

Here, we used agglomerative hierarchical clustering (Johnson, 1967), a simple, intuitive and well-established method. This method proceeds with repeated steps of calculating distances between data points, merging 'closest' points into one cluster, and re-calculating distances until all the data are in one cluster. We used Euclidean distance (sum of squared differences) as our measure of similarity, a measure of 'closeness' which groups participants' data based on the magnitude of the features. Since we first standardize our features to account for any differences in scales of measurement, Euclidean distance is a valid measure of similarity, and given its simplicity, a straightforward choice. Mergers are done such that the overall increase in within-cluster variance at each step is kept to a minimum, known as the Ward Linkage Method (Ward, 1963) for hierarchical clustering. One advantage of this is approach is that the output is in the form of a tree-like structure known as a dendrogram, which shows the different degrees of similarity where data points are merged into groups. The dendrogram is used to visualize the number of clusters in the data based on where larger differences in Euclidean distance are observed as the data points are clustered into smaller and smaller groups.

Clustering analyses were conducted using an in-house code written in Python 2.7 (www.python.org), specifically making use of scikit-learn (Pedregosa et al., 2011), and scipy (Jones, Oliphant, & Peterson, 2014) libraries. Input data to the unsupervised clustering algorithm were N (i.e., number of subjects) feature vectors of length T*K where T is the number of trials and K is the number of physiological variables. This clustering approach adopts a 2-level data structure, nesting measures and trials within participants, and so controls for participant random effects. Each participant was assigned a feature vector that included as features that participant's physiological reactivity (i.e., difference scores between task and baseline activity) for all trials of the task for all recorded measures. These individual feature vectors (one for each participant) were then clustered, allowing us to examine across-participant similarity (or dissimilarity) in both (1) the within-participant pattern of mean physiological reactivity across measures and (2) the within-participant pattern of change in physiological reactivity across trials. To remove the effects of different measurement scales for each variable, we first standardized each feature across the whole dataset by subtracting the mean and dividing by its standard deviation. These standardized feature vectors became the input to the hierarchical clustering algorithm, which gives the dendrogram as its output.

Clustering results were then obtained by deciding where to cut the dendrogram by visual inspection. We also calculated silhouette scores as a measure of goodness-of-fit (using Python's sci-kit learn library implementation) for clustering solutions with 2 to 6 clusters for each dataset (See Table 2). Visual inspection and the silhouette scores both pointed to viable 2- or 3-cluster solutions with poorer fit for clustering solutions with a greater number of clusters, particularly for Dataset 2. For completeness, we examined both 2- and 3-cluster solutions to determine which had the more consistent and interpretable solution across datasets.

3.2 Feature Relevance

Once clustering resulted in C clusters, each of the individual physiological features was ranked by measuring its 'contribution' to the clustering result. We utilized two measures for this purpose. One measure was a *scatter ratio*, which assesses the ratio of between-cluster variance to within-cluster variance for a given feature. This can be evaluated for each feature as an indicator of how well separated the clusters are in that feature's space. A higher scatter ratio indicates that the feature varies much more dramatically across clusters than within-clusters (i.e., more well-separated but compact clusters (Wang, Li, Song, Wei, & Li, 2011)). For a feature f_k, the scatter-ratio (J_{fk}) is given by the between-cluster scatter divided by the within-cluster scatter. Formulas for calculating between- and within-cluster scatter can be found in the supplemental materials online.

The second measure of feature relevance we employed was mutual information (*MI*). *MI* can be interpreted as how much the knowledge of one random variable (here, one feature) reduces the uncertainty about another random variable (here, the cluster labels). A higher value indicates a higher reduction in uncertainty, thus more informative features will exhibit a higher *MI* value. Because *MI* is unbounded, we standardized this measure so that values range between 0 and 1 (also called the global correlation coefficient (Darbellay & Wuertz, 2000) as follows:

$$\lambda_k = \left(1 - \exp\left(-2I\left(\boldsymbol{x_k}, \boldsymbol{y}\right)\right)^{\frac{1}{2}}\right.$$

where, $I(\boldsymbol{x_k}, \boldsymbol{y})$ is the mutual information between feature f_k and labels.

3.3 Self-report Appraisals of Threat and Challenge

We calculated stress-coping ratios for both datasets as a way to measure participants' self-reported appraisals of threat and challenge. For Dataset 1, appraisals were assessed via two multiple-choice items. Participants rated: (1) how stressful they perceived the task to be and (2) how well they thought they would (or did) cope with the task on two 5-point scales (from 1 = "Not at All" to 5 = "Very Much"). For Dataset 1, we divided participants' self-reported stress ratings by their self-reported coping resources ratings for each of the three appraisal ratings collected throughout the mental math task (i.e., before, during, and after the mental math task). For Dataset 2, appraisals were assessed via two self-report items. Participants rated: (1) how stressful they perceived the task to be on a 9-point scale (from −4 = "strongly disagree" to 4 = "strongly agree") and (2) their perceived resources for dealing with the task on the same 9-point scale. Responses to both questions were re-scored from 1 = "strongly disagree" to 9 = "strongly agree" for analyses. For Dataset 2, we calculated a stress/coping ratio by dividing participants' self-reported stress ratings by their self-reported resources for each of the two sets of appraisal ratings (i.e., before and after the mental math task). We assessed these stress-coping ratios after clustering results were known to examine whether any physiologically-based clusters differed in terms of their associated appraisals of threat and challenge. In particular, we examined whether clusters differed in the extent to which members of the cluster were likely to self-report experiencing challenge (i.e., where their coping resources met or exceeded the perceived stressfulness of the task, such that stress-coping ratios were ≤1) versus threat (i.e., where the perceived stressfulness of the task exceeded their coping resources, such that stress-coping ratios were >1). For Dataset 1, stress and coping ratings were missing for several participants due to a computer programming error at the start of data collection; appraisal ratings are thus reported from the $n = 150$ participants with appraisal data.

4 Results

Dataset 1 consisted of 165 participants with 12 physiological variables (MAP, CO, PEP, TPR, LVET, SCRs, SCL, CORR, ZYGO, RSA, IBI, and RR) across 4, one-minute trials of a mental math task. Dataset 2 consisted of 130 participants with 10 physiological variables (same variables as Dataset 1, but without CORR and ZYGO) across 5, one-minute trials of a mental math task. Basic descriptive statistics and reliability estimates for each reactivity measure for each dataset are provided in Table 3.

We used hierarchical clustering on our two datasets, and visual inspection of the clustering dendrograms (Figures 1 and 2) revealed several possible cut-off points. Although a two-cluster solution is evident for both datasets, Dataset 2 also appears to offer a viable 3-cluster solution (Figure 2). Consistent with this visual inspection, silhouette scores calculated as measures of goodness-of-fit for the 2- and 3-cluster solutions for both datasets (see Table 2) were fairly comparable. For completeness, we compared both 2-cluster and 3-cluster solutions across the two datasets. However, we report only the detailed results for the 3-cluster solution in the main text, with detailed results for the 2-cluster solution reported in the supplemental online materials. After comparing both 2-cluster and 3-cluster solutions, we chose the 3-cluster solution as superior to the 2-cluster solution because of the consistency of the patterns of physiological reactivity across clusters in the 3-cluster solutions across the two independent datasets.

4.1 Two-Cluster Solution

Dendrograms for both datasets (Figures 1 and 2) included cut-off points that resulted in 2 clusters (above 28 for Dataset 1 and above 32 for Dataset 2). These 2-cluster solutions both resulted in one larger cluster (104 members in Dataset 1; 108 members in Dataset 2) and one smaller cluster. In both datasets, the two clusters in the 2-cluster solutions appeared to map well on to the patterns of physiological activity predicted by the existing literature on challenge and threat states (see Figures S1 and S3): one cluster in each dataset revealed a challenge-like pattern of physiological response (i.e., increased CO coupled with decreased TPR and IBI), whereas the other cluster in each dataset revealed a threat-like pattern of physiological response (i.e., minimal changes in CO, TPR, and IBI). These clusters corresponded with predicted patterns in self-reported stress/coping ratios for ratings made prior to the start of the mental math task, although this pattern was less pronounced or not present for ratings made during or after the mental math task (see Figures S2 and S4). Feature relevance scores for the 2-cluster solutions revealed some parallels across datasets (e.g., CO and IBI were significant contributors to clustering solutions in both datasets) as well as several inconsistencies across datasets; of note, SCL was the most significant feature contributing to the 2-cluster solution for Dataset 2, but was the second least most significant feature in Dataset 1 (see Table S1). For more detailed discussion of the 2-cluster solutions, see the supplemental online materials.

4.2 Three-Cluster Solution

Dendrograms for both datasets also included cut-off points that resulted in 3 clusters (above 25 for Dataset 1 and above 26 for Dataset 2). The 3-cluster solution for Dataset 1 resulted in one cluster with 33 members (shown in black), one cluster with 71 members (shown in blue), and one cluster with 61 members (shown in green); compared to the 2-cluster solution, the 3-cluster solution for Dataset 1 involved the larger cluster from the 2-cluster solution being split into two smaller clusters (here black and blue) while the other, smaller cluster (here green) was unchanged. The 3-cluster solution for Dataset 2 resulted in one cluster with 9 members (shown in black), one cluster with 13 members (shown in green), and one cluster with 108 members (shown in blue); compared to the 2-cluster solution, the 3-cluster solution for Dataset 2 involved the smaller cluster from the 2-cluster solution being split into two smaller clusters (here black and green) while the other, larger cluster (here blue) was unchanged.

4.2.1 Dataset 1.—In the 3-cluster solution for Dataset 1 (Figure 3), the green cluster appeared to have a challenge-like physiological pattern of response with increased CO accompanied by decreased TPR, PEP, and IBI. Conversely, the blue and black clusters both mostly exhibited a threat-like physiological pattern of response with minimal changes in CO, TPR, PEP, and IBI (with the blue cluster exhibiting the most minimal changes in these variables). However, the blue and black clusters were most clearly differentiated by changes in SCL, with individuals within the black cluster exhibiting pronounced increases in SCL and individuals within the blue cluster exhibiting only minimal changes in SCL. These patterns were consistent with an interpretation of the black cluster as being comprised of individuals exhibiting a threat-like pattern of physiological response and the blue cluster as being comprised of physiological non-responders, quite possibly those who did not find the task stressful and so did not respond as if it were an active coping stressor task.

The pattern of self-reported appraisals of threat and challenge from before the start of the mental math task appeared consistent with this interpretation (Figure 4a). As in the two-cluster solution (see supplemental online materials), a larger proportion of individuals from the green “challenge” cluster reported that they expected their coping resources would meet or exceed the stressfulness of the task (i.e., a challenge appraisal) compared to those in black “threat cluster”, and a larger proportion of individuals in the black “threat” cluster reported that their stress exceeded their coping resources (i.e., a threat appraisal) compared to those the green “challenge” cluster ($X^2(1, N = 89) = 6.43$, $p = .01$). The blue “non-responder” cluster had intermediate proportions of individuals reporting threat and challenge appraisals, which did not differ significantly from proportions in either the green “challenge” cluster ($X^2(1, N = 120) = 2.32$, $p = .12$) or the black “threat” cluster ($X^2(1, N = 93) = 1.55$, $p = .21$). However, this pattern did not hold for self-reported appraisals collected *during* the mental math task (Figure 4b; black v. green clusters: $X^2(1, N = 89) = 0.46$, $p = .49$; black v. blue clusters: $X^2(1, N = 93) = 1.47$, $p = .22$; blue v. green clusters: $X^2(1, N = 120) = 0.41$, $p = .52$), and was only weakly (and non-significantly) observed in appraisals reported *after* the mental math task (Figure 4c; black v. green clusters: $X^2(1, N = 89) = 0.40$, $p = .52$; black v. blue clusters; $X^2(1, N = 93) = 0.20$, $p = .66$); green v. blue clusters: $X^2(1, N = 120) = 0.05$, $p = .81$).

4.2.2 Dataset 2.—Consistent with the results from Dataset 1, the results for the 3-cluster solution revealed a green "challenge" cluster evidenced by pronounced increases in CO accompanied by pronounced decreases in TPR, PEP, and IBI; a black "threat" cluster with minimal changes in CO, TPR, PEP, and IBI but large increases in SCL, and a blue "non-responder" cluster with minimal changes across all physiological variables (Figure 5). However, self-reported appraisals of threat and challenge from before (Figure 6a) and after (Figure 6b) the mental math task were not entirely consistent with this interpretation. For instance, individuals from the black "threat" group all reported that they expected their coping resources would meet or exceed the stressfulness of the task (a 'challenge' appraisal) prior to the start of the mental math task. Nevertheless, this pattern of self-reported appraisals should be interpreted with caution because inferential statistics were not performed on self-reported appraisals for Dataset 2 because of the dramatically uneven cluster sizes and because at least one comparison condition contained five or fewer individuals.

4.2.3 Feature Relevance Scores.—Feature relevance scores describing the contribution for each physiological feature towards the 3-cluster solutions for both Dataset 1 and Dataset 2 are given in Table 4. For Dataset 1 the physiological features most relevant to the clustering solution are CO, TPR, IBI, and SCL (followed by PEP and LVET) across both feature relevance metrics. Consistent with this, in Dataset 2, CO, IBI, SCL, and PEP are the top four most relevant features across both feature relevance metrics. However, as in the two-cluster solution (see supplemental online materials), TPR is only the 5^{th} or 6^{th} most relevant feature out of the 10 features utilized in Dataset 2. LVET is the least relevant feature for the three-cluster solution in Dataset 2. The primary difference in feature relevance rankings between the 2- and 3-cluster solutions is that SCL is among the top-most relevant features across both data sets in the 3-cluster solutions.

4.2.4 Baseline Differences across Clusters.—For the 3-cluster solutions, we also examined differences across clusters in basic demographic and anthropometric measures as well as baseline physiological activity (see Table 5). For Dataset 1, a number of baseline and demographic variables differed significantly across clusters (see Table 5). The blue 'non-responder' cluster in Dataset 1 was significantly older than the other two clusters and contained a significantly smaller proportion of White participants than did the green challenge-like cluster. The blue 'non-responder' cluster in Dataset 1 also had lower baseline LVET, RR, RSA, and TPR than the other clusters. In Dataset 1, the green challenge-like cluster had significantly lower baseline CO than did either of the other two clusters. For Dataset 2, only SCL differed significantly across clusters: SCL was highest in the black threat-like cluster and lowest in the blue 'non-responder' cluster with intermediate SCL in the green challenge-like cluster. However, comparisons across clusters within Dataset 2 must be interpreted with caution given the uneven cluster sizes and the inclusion of two relatively small clusters ($n = 9$ and $n = 13$).

4.2.5 Performance and Appraisal Ratings Across Clusters.—Table 6 reports mean stress appraisals and mean coping appraisals individually across clusters as well as two metrics of task performance (i.e., responses attempted and proportion of correct responses). We see no significant differences across clusters on any of these metrics across either sample. In line with previous empirical work, the ratio of stress to coping resources appears to be more strongly associated with patterns of peripheral physiological activity than do either the stress or coping appraisals individually. Also, while one might anticipate performance to be better on average among individuals in the challenge-like clusters than the threat-like clusters, that does not appear to be the case in the present samples. We suspect this is because in both studies the difficulty of the serial subtractions was adjusted throughout the course of the task based on performance in order to keep the task consistently stressful for participants.

5 Discussion

Using a data-driven, unsupervised machine learning approach, we identified two to three groups of individuals who had similar patterns of physiological responding during an active coping stressor task, and we replicated these groupings across two large, independent data sets. To the best of our knowledge, this is the first study to attempt to cluster individuals based solely on their physiological activity during a motivated performance task, without the use of self-reported appraisal scores or subject variables (e.g., age, race, gender) to guide the clustering solution. Thus, it provided a uniquely strong test of whether physiological patterns of activity commonly associated with the psychobiological states of threat and challenge are actually the *predominant* categories of physiological response in such contexts. Our results are strikingly consistent with this interpretation, although the results also highlight a third predominant pattern, of non-responding indicated by minimal change from baseline across multiple measures of peripheral physiological activity.

In the three-cluster solutions for the two datasets, we saw two clusters of individuals emerging in both datasets, one who appeared to have challenge-like patterns of physiological activity (i.e., increased CO and decreased PEP, IBI, and TPR) and one with threat-like patterns of physiological activity (i.e., more modest change in these same variables). As in the two-cluster solutions (see supplemental online materials), measures of left ventricular contractility were also highly relevant to cluster differentiation, although not consistently so across datasets (e.g., PEP, particularly in Dataset 2, and LVET in Dataset 1 only). In addition, however, we characterized a third cluster of individuals who showed minimal change in physiological activity relative to baseline across all measures. Although this cluster of physiological ‘non-responders’ had more modest changes from baseline (if any) compared to the threat-like cluster across the standard cardiovascular indices of threat and challenge, the non-responders were best differentiated from the other two groups by a lack of change in tonic sweat gland activity (i.e., SCL). Consistent with theorizing, both the challenge-like and threat-like groups demonstrated increased sympathetic nervous system activation (here, indicated by increases in SCL). Increased SCL in the threat-like group is also consistent with previous work demonstrating greater electrodermal activity in response to a wide variety of aversive stimuli or potential threats, including threat of electric shock (see, e.g., Niemelá, 1969; Kopacz & Smith, 1971). In contrast, the ‘non-responder’ group showed minimal change in SCL activity relative to baseline across both datasets. In the two datasets examined here, SCL has clear discriminable power, as was also shown in other recent work (Nagai et al., 2004), despite the fact that much recent work in psychophysiology focuses on skin conductance responses, rather than on SCL. Thus, despite its minimal use in this literature, SCL joins measures of cardiac function (i.e., CO and IBI) and vascular resistance (i.e., TPR) as one of the most relevant features for differentiating a 3-cluster solution across both data sets and via both feature relevance metrics. Together with the other measures, it provides a useful way to distinguish 'non-responders' from those exhibiting either threat or challenge patterns of responding.

Of note, in Dataset 2, the cluster of individuals we originally identified as having a threat-like pattern of physiological responding in the 2-cluster solution (see supplemental online materials) was re-categorized and re-interpreted in the 3-cluster solution as a group of physiological non-responders. Because the physiological pattern of response for threat is typically defined in relation to that of a challenge pattern (i.e., as more modest change in cardiac output, ventricular contractility, and peripheral resistance compared to challenge), it was straight-forward and consistent with the literature to identify the larger cluster in the two-cluster solution as indicative of a threat state. Yet, when we considered the possibility of a third cluster, it became apparent that there was a more appropriate interpretation, namely that these individuals exhibited minimal change in physiological activity relative to baseline across all physiological measures to the active coping stressor task. The high feature relevance scores for SCL for this dataset in the 2-cluster solution is also consistent with the idea that the 'threat' cluster for the 2-cluster solution for Dataset 2 was actually a misinterpreted 'non-responder' cluster. This re-interpretation of the same cluster of individuals highlights how the focus of the existing literature on two orthogonal psychobiological states, to the exclusion of possible others, may artificially constrain interpretations of observed patterns of ANS activity within motivated performance contexts. This also may help to explain why the biological pattern of responding under threat in particular has been somewhat inconsistent across studies and tasks. For example, because threat and challenge are typically identified within a study only in relation to one another, threat has been characterized as involving 'minimal or 'no change' in CO and an 'increase' or 'no change' in TPR (e.g., Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996; Mendes et al., 2002; Quigley et al., 2002; Tomaka et al., 1997). Our findings suggest that some of this inconsistency in the pattern of observed ANS activity under threat may be due to differences in the prevalence of 'non-responders' and their distribution across the threat and challenge groups across contexts and samples.

The identification of a group of apparent physiological 'non-responders' in both our datasets is consistent with the results of a recent study which utilized a similar approach (multivariate cluster analysis with a small set of variables) to identify groups of individuals based on their physiological reactivity during a set of psychological stress tasks (Brindle et al., 2016). Although the study only included three cardiovascular measures—heart rate, systolic blood pressure, and diastolic blood pressure—their data-driven cluster analysis also revealed a group of participants who exhibited blunted physiological reactivity during motivated performance tasks across all recorded measures. Given this consistency, future research may benefit from exploring the extent to which these patterns of physiological response represent more stable individual differences or responder types across contexts. For example, prior research has demonstrated responder type differences in cardiovascular reactivity and recovery during speech preparation and cold presser stress tasks (Kline et al., 2002). Future research should address the stability of all of these physiological patterns across time within individuals, given prior evidence that repeated intense vascular reactivity and/or slower vascular recovery confers negative cardiovascular health risk (Sherwood et al., 1999; Treiber et al., 1993; Treiber et al., 2003).

Our results also highlight several practical methodological suggestions for researchers interested in examining threat and challenge states and/or ANS function in motivated performance contexts. First, it appears unnecessary to include several peripheral physiological measures tested here in future investigations using a motivated mental math task. For example, measures of respiration rate and heart rate variability driven by respiration (i.e., RSA) as well as measures of facial muscle activity over the corrugator supercilii and zygomaticus major muscle regions of the face and the number of skin conductance responses were all routinely among the least relevant physiological features for distinguishing between clusters, in both the 2- and 3-cluster solutions, across both independent datasets. However, given that both datasets were collected using the same active coping stressor task (i.e., mental math), future such data-driven work with other tasks should be used to determine if the same holds across other motivated performance contexts. Second, despite few prior studies of the challenge v. threat patterns that used any measures of skin conductance (only Experiments 1 & 3 of Tomaka et al., 1993, to our knowledge), our findings strongly advocate for the inclusion of measures of SCL, which may be of particular importance when researchers are trying to distinguish threat-like responding from physiological non-response. Finally, our findings suggested that self-reported appraisals of stress and coping made *prior to* the start of the task were more meaningfully associated with patterns of physiological responding than were appraisals made after the task, a finding that is empirically and theoretically consistent with the preponderance of the existing literature (see, e.g., Quigley et al., 2002; Tomaka et al., 1997; Zanstra, Johnston, & Rasbash, 2010). Nevertheless, even the differences we observed in reported appraisals across clusters before the start of the task were modest, suggesting the need for further refinement of ways to assess perceived stress and coping in motivated performance tasks.

5.1 Limitations and Future Directions

The present investigation has several limitations that suggest fruitful avenues for future research. First, both datasets also appear to have a large proportion of physiological ‘non-responders’ (43% in Dataset 1 and 83% in Dataset 2), with this cluster particularly prominent in Dataset 2. One possibility is that the prevalence of non-responders in the current datasets is an artifact of our exclusion criteria for analyses. In the present investigation, our analysis strategy necessitated limiting our sample to only those individuals with ‘complete’ data sets (i.e., those with usable physiological data from each minute of the task and baseline for all 12 (Dataset 1) or 10 (Dataset 2) physiological features). This criterion may have created a biased sample selection favoring the inclusion of non-responders in the final sample used for analysis (i.e., more distressed people may fidget and produce more movement artifacts and do so more often). However, the difference in the proportion of non-responders across datasets is also striking and suggests that elements of the two experimental designs may have contributed to the varying occurrence of this pattern of physiological response across datasets. Of note, in Dataset 2, participants completed a different motivated performance task, a speech preparation and delivery task, prior to completing the mental math task analyzed herein. Previous research has demonstrated that there are significant patterns of cardiovascular adaptation to repeated active coping tasks (e.g., mental arithmetic), and that cardiac reactivity is especially sensitive to prior exposure to another motivated performance task (Kelsey, Blascovich, Tomaka, Leitten, Schneider, & Wiens, 1999; Kelsey, Blascovich, Leitten, Schneider, Tomaka & Wiens, 2000; Kelsey, Soderlund, & Arthur, 2004; Kelsey, Ornduff, & Alpert, 2007). As such, this pattern of physiological 'non-response' may be associated with or even caused by adaptation to repeated psychological stressor task performance. Future research should compare the prevalence of these three patterns of psychobiological response across different active coping stressor tasks, as well as across experimental contexts that vary in terms of the extent of prior exposure to additional motivated performance tasks.

The identification of a third pattern of physiological activity during an active coping stressor task also lays bare several interesting potential future lines of inquiry concerning how this pattern of ANS activity relates to both task performance and subjective appraisals of experience. That is, although we have interpreted this third group as 'non-responders', this label refers solely to their absent or minimal physiological response relative to baseline and is only one of a host of possible reasons for this pattern of response which should be explored in future research. Indeed, the pattern of self-reported appraisals across the three clusters of physiological activity differs across the two independent datasets in the present study. For pre-task appraisal ratings in Dataset 1, a larger proportion of individuals from the threat-like cluster reported that they expected that their coping resources would not meet the perceived demands of the task (i.e., a threat appraisal) relative to both the challenge-like and non-responder clusters; although a sizable proportion of the individuals in all clusters, including the threat-like cluster, self-reported appraisals that were closer to the challenge end of the challenge-threat spectrum (i.e., appraisals of task stressfulness were somewhat mitigated by appraisals that the person felt they could cope with the demands of the task). This pattern is consistent with the possibility that the physiological non-responders may also be 'non-responders' from a psychological perspective, such that they may simply not be experiencing the task as an active coping stressor task (i.e., they perceive the task demands as very low or their own relative resources for coping with the task as very high). However, this pattern failed to emerge for appraisal ratings in Dataset 2 where a smaller proportion of 'non-responders' reported challenge-like appraisals than even those in the threat-like cluster. Thus, the cluster of physiological non-responders may not consistently be associated with the same pattern of self-reported appraisals across samples. It appears that, at least in some contexts, the lack of physiological activity relative to baseline exhibited by these individuals may be independent of their experience of the stressor task, such that some or all of these physiological non-responders still report experiencing a great deal of distress in anticipation of the start of the task. Indeed, we found no differences across clusters in either dataset in terms of average self-reported stress or coping appraisals when these appraisals were examined independently, and we also failed to find any differences across clusters in either dataset in the number of responses attempted or the proportion of correct responses. These findings are inconsistent with an interpretation of the 'non-responder' cluster as disengaged or non-stressed. Nevertheless, we interpret these null results with caution, particularly for

Dataset 2 given the dramatically uneven size of the clusters in this clustering solution (i.e., the threat-like, challenge-like, and non-responder clusters for Dataset 2 contain 9, 13, and 108 participants, respectively). We encourage future researchers to investigate how appraisals of threat and challenge map on to groups of individuals identified as physiological non-responders. Moreover, given the differences across datasets in patterns of self-reported appraisals, future research should examine whether exposure to prior motivational performance tasks and associated adaptations in cardiovascular responding (see Kelsey et al., 1999; 2000; 2004; 2007) may contribute to differential relationships between physiological 'non-response' and self-reported appraisals of challenge and threat.

5.2 Conclusions

We examined patterns of peripheral physiological responding during a motivated performance context across two large, independent data sets, each with multiple peripheral physiological measures. A data-driven, unsupervised machine learning approach revealed interpretable 3-cluster solutions across both datasets with consistent patterns of physiological responding across clusters in both independent samples. Two of the predominant patterns of peripheral physiological responding that emerged within both samples were strikingly similar to cardiovascular responses commonly associated with 'challenge' and 'threat' states, with these two patterns were best differentiated by reactivity in cardiac output (CO), pre-ejection period (PEP), interbeat interval (IBI), and total peripheral resistance (TPR). Results also revealed a third, relatively large group of physiological "non-responders" who exhibited minimal reactivity across all physiological measures in the motivated performance context, and were most differentiated from those in the other clusters by minimal increases in tonic electrodermal activity. Future research is needed to more fully characterize this group of physiological 'non-responders' including its relationship with individual differences in physiological responding across a wider range of evocative contexts, its relationship with prior exposure to other psychological stressors, and its relationship to subjective experiences of stress.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

This research was supported by the U.S. Army Research Institute for the Behavioral and Social Sciences (W5J9CQ-12-C-0049 to L.F.B., W911N-16–1-0191 to K.S.Q. and J.B.W., and W5J9CQ-12-C-0028 to S.K.L) and by the National Institutes of Health (R01MH113234 and 1U01CA193632–01A1 to L.F.B.). The views, opinions, and/or findings contained in this paper are those of the authors and shall not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documents.

References

Alter AL, Aronson J, Darley JM, Rodriguez C, & Ruble DN (2010). Rising to the threat: Reducing stereotype threat by reframing the threat as a challenge. Journal of Experimental Social Psychology, 46(1), 166–171. 10.1016/j.jesp.2009.09.014

Barbakh WA, Wu Y, & Fyfe C (2009). Non-standard parameter adaptation for exploratory data analysis. Berlin: Springer.

Barrett LF (2017). How emotions are made: The secret life of the brain. New York, NY: Houghton-Mifflin-Harcourt

Barrett LF (2017). The theory of constructed emotion: an active inference account of interoception and categorization. Social Cognitive and Affective Neuroscience, 12(1), 1–23. 10.1093/scan/nsw154 [PubMed: 27798257]

Blascovich J, & Mendes W (2000). Challenge and threat appraisals: The role of affective cues In Forgas J (Ed.), Feeling and thinking: The role of affect in social cognition (pp. 59–82). Cambridge, UK: : Cambridge University Press.

Blascovich J, Mendes WB, Hunter SB, Lickel B, & Kowai-Bell N (2001). Perceiver threat in social interactions with stigmatized others. Journal of Personality and Social Psychology, 80(2), 253–267. 10.1037/0022-3514.80.2.253 [PubMed: 11220444]

Blascovich J, & Tomaka J (1996). The biopsychosocial model of arousal regulation. Advances in Experimental Social Psychology, 28, 1–51. 10.1016/S0065-2601(08)60235-X

Brindle RC, Ginty AT, Jones A, Phillips AC, Roseboom TJ, Carroll D, … & de Rooij, S. R. (2016). Cardiovascular reactivity patterns and pathways to hypertension: a multivariate cluster analysis. Journal of human hypertension, 30(12), 755–760. 10.1038/jhh.2016.35 [PubMed: 27334523]

Brimmell J, Parker JK, Furley P, & Moore LJ (2018). Nonverbal behavior accompanying challenge and threat states under pressure. Psychology of Sport and Exercise, 39, 90–94. 10.1016/j.psychsport.2018.08.003

Chanes L, & Barrett LF (2016). Redefining the role of limbic areas in cortical processing. Trends in Cognitive Sciences, 20(2), 96–106. 10.1016/j.tics.2015.11.005 [PubMed: 26704857]

Darbellay GA, & Wuertz D (2000). The entropy as a tool for analysing statistical dependences in financial time series. Physica A: Statistical Mechanics and its Applications, 287(3), 429–439. 10.1016/S0378-4371(00)00382-4

Dienstbier RA (1989). Arousal and physiological toughness: implications for mental and physical health. Psychological Review, 96(1), 84–100. 10.1037/0033-295X.96.1.84 [PubMed: 2538855]

Drach-Zahavy A, & Erez M (2002). Challenge versus threat effects on the goal–performance relationship. Organizational Behavior and Human Decision Processes, 88(2), 667–682. 10.1016/S0749-5978(02)00004-3

Harvey A, Nathens AB, Bandiera G, & LeBlanc VR (2010). Threat and challenge: cognitive appraisal and stress responses in simulated trauma resuscitations. Medical Education, 44(6), 587–594. 10.1111/j.1365-2923.2010.03634.x [PubMed: 20604855]

Jamieson JP, Mendes WB, Blackstock E, & Schmader T (2010). Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. Journal of Experimental Social Psychology, 46(1), 208–212. 10.1016/j.jesp.2009.08.015 [PubMed: 20161454]

Johnson SC (1967). Hierarchical clustering schemes. Psychometrika, 32(3), 241–254. 10.1007/BF02289588 [PubMed: 5234703]

Jones E, Oliphant T, & Peterson P (2014). {SciPy}: Open source scientific tools for {Python}. Retrieved from https://www.scipy.org/

Kelsey RM, Blascovich J, Leitten CL, Schneider TR, Tomaka J, & Wiens S (2000). Cardiovascular reactivity and adaptation to recurrent psychological stress: The moderating effects of evaluative observation. Psychophysiology, 37(6), 748–756. 10.1111/1469-8986.3760748 [PubMed: 11117455]

Kelsey RM, Blascovich J, Tomaka J, Leitten CL, Schneider TR, & Wiens S (1999). Cardiovascular reactivity and adaptation to recurrent psychological stress: Effects of prior task exposure. Psychophysiology, 36(6), 818–831. 10.1111/1469-8986.3660818 [PubMed: 10554594]

Kelsey RM, Ornduff SR, & Alpert BS (2007). Reliability of cardiovascular reactivity to stress: Internal consistency. Psychophysiology, 44(2), 216–225. 10.1111/j.1469-8986.2007.00499.x [PubMed: 17343705]

Kelsey RM, Soderlund K, & Arthur CM (2004). Cardiovascular reactivity and adaptation to recurrent psychological stress: Replication and extension. Psychophysiology, 41(6), 924–934. 10.1111/j.1469-8986.2004.00245.x [PubMed: 15563345]

Kline KA, Saab PG, Llabre MM, Spitzer SB, Evans JD, McDonald PAG, & Schneiderman N (2002). Hemodynamic response patterns: Responder type differences in reactivity and recovery. Psychophysiology, 39(6), 739–746. 10.1111/1469-8986.3960739 [PubMed: 12462502]

Kopacz FM, & Smith BD (1971). Sex differences in skin conductance measures as a function of shock threat. Psychophysiology, 8(3), 293–303. 10.1111/j.1469-8986.1971.tb00459.x [PubMed: 5093972]

Mendes WB, Blascovich J, Hunter SB, Lickel B, & Jost JT (2007). Threatened by the unexpected: physiological responses during social interactions with expectancy-violating partners. Journal of Personality and Social Psychology, 92(4), 698–716. 10.1037/0022-3514.92.4.698 [PubMed: 17469953]

Mendes WB, Blascovich J, Lickel B, & Hunter S (2002). Challenge and threat during social interactions with White and Black men. Personality and Social Psychology Bulletin, 28(7), 939–952. 10.1177/01467202028007007

Mendes WB, Blascovich J, Major B, & Seery M (2001). Challenge and threat responses during downward and upward social comparisons. European Journal of Social Psychology, 31, 477–497. 10.1002/ejsp.80

Mendes WB, Gray H, Mendoza-Denton R, Major B, & Epel E (2007). Why egalitarianism might be good for your health: Physiological thriving during inter-racial interactions. Psychological Science, 18, 991–998. 10.1111/j.1467-9280.2007.02014.x [PubMed: 17958714]

Nagai Y, Critchley HD, Featherstone E, Trimble MR, & Dolan RJ (2004). Activity in ventromedial prefrontal cortex covaries with sympathetic skin conductance level: A physiological account of a “default mode” of brain function. Neuroimage, 22(1), 243–251. 10.1016/j.neuroimage.2004.01.019 [PubMed: 15110014]

Niemelá P (1969). Electrodermal responses as a function of quantified threat. Scandinavian Journal of Psychology, 10(1), 49–56. 10.1111/j.1467-9450.1969.tb00007.x [PubMed: 5353398]

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, … & Vanderplas J (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825–2830.

Qu M, Zhang Y, Webster JG, & Tompkins WJ (1986). Motion artifact from spot and band electrodes during impedance cardiography. IEEE Transactions on Biomedical Engineering (11), 1029–1036. 10.1109/TBME.1986.325869 [PubMed: 3793123]

Quigley KS, Barrett LF, & Weinstein S (2002). Cardiovascular patterns associated with threat and challenge appraisals: A within-subjects analysis. Psychophysiology, 39(3), 292–302. 10.1017/S0048577201393046 [PubMed: 12212648]

Seery MD (2011). Challenge or threat? Cardiovascular indexes of resilience and vulnerability to potential stress in humans. Neuroscience & Biobehavioral Reviews, 35(7), 1603–1610. 10.1016/j.neubiorev.2011.03.003 [PubMed: 21396399]

Skinner N, & Brewer N (2002). The dynamics of threat and challenge appraisals prior to stressful achievement events. Journal of Personality and Social Psychology, 83(3), 678–692. 10.1037/0022-3514.83.3.678 [PubMed: 12219862]

Sherwood A, Johnson K, Blumenthal JA, & Hinderliter AL (1999). Endothelial function and hemodynamic responses during mental stress. Psychosomatic Medicine, 61(3), 365–370. 10.1097/00006842-199905000-00017 [PubMed: 10367618]

Streamer L, Seery MD, Kondrak CL, Lamarche VM, & Saltsman TL (2017). Not I, but she: The beneficial effects of self-distancing on challenge/threat cardiovascular responses. Journal of Experimental Social Psychology, 70, 235–241. 10.1016/j.jesp.2016.11.008

Tomaka J, Blascovich J, Kelsey RM, & Leitten CL (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. Journal of Personality and Social Psychology, 65(2), 248–260. 10.1037/0022-3514.65.2.248

Tomaka J, Blascovich J, Kibler J, & Ernst JM (1997). Cognitive and physiological antecedents of threat and challenge appraisal. Journal of Personality and Social Psychology, 73(1), 63–72. 10.1037/0022-3514.73.1.63 [PubMed: 9216079]

Treiber FA, Davis H, Musante L, Raunikar RA, Strong WB, McCaffrey F, … & Vandernoord R(1993). Ethnicity, gender, family history of myocardial infarction, and hemodynamic responses to laboratory stressors in children. Health Psychology, 12(1), 6–15. 10.1037/0278-6133.12.1.6 [PubMed: 8462501]

Treiber FA, Kamarck T, Schneiderman N, Sheffield D, Kapuku G, & Taylor T (2003). Cardiovascular reactivity and development of preclinical and clinical disease states. Psychosomatic Medicine, 65(1), 46–62. 10.1097/00006842-200301000-00007 [PubMed: 12554815]

Wang S, Li D, Song X, Wei Y, & Li H (2011). A feature selection method based on improved fisher's discriminant ratio for text sentiment classification. Expert Systems with Applications, 38(7), 8696–8702. 10.1016/j.eswa.2011.01.077

Ward JH Jr (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58(301), 236–244. 10.1080/01621459.1963.10500845

Zanstra YJ, Johnston DW, & Rasbash J (2010). Appraisal predicts hemodynamic reactivity in a naturalistic stressor. International Journal of Psychophysiology, 77(1), 35–42. 10.1016/j.ijpsycho.2010.04.004 [PubMed: 20417669]

Zilka GC, Rahimi ID, & Cohen R (2019). Sense of challenge, threat, self-efficacy, and motivation of students learning in virtual and blended courses. American Journal of Distance Education, 33(1), 2–15. 10.1080/08923647.2019.1554990

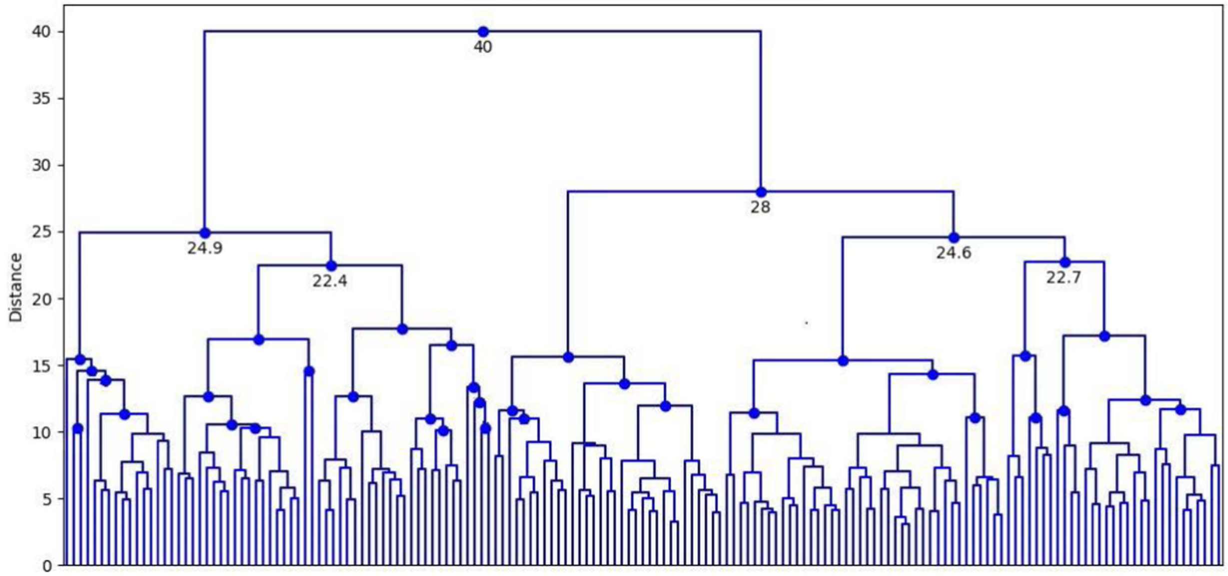

Figure 1.
Dendrogram for Dataset 1with Euclidean distance on the y-axis and participant number on the x-axis.

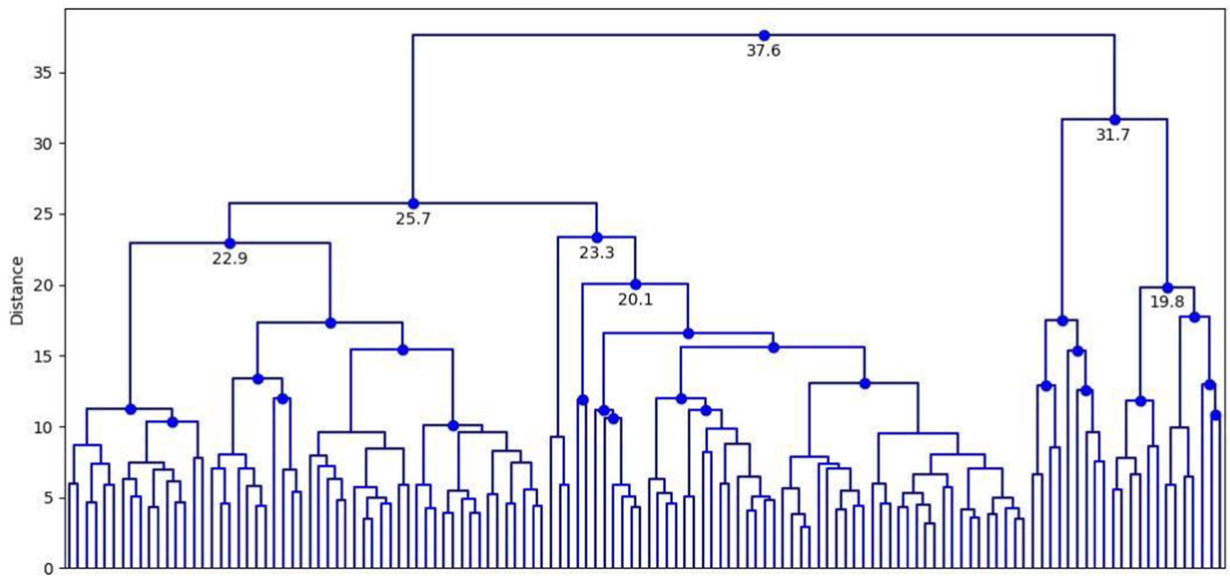

Figure 2.
Dendrogram for Dataset 2 with Euclidean distance on the y-axis and participant number on the x-axis.

Author Manuscript
Author Manuscript
Author Manuscript

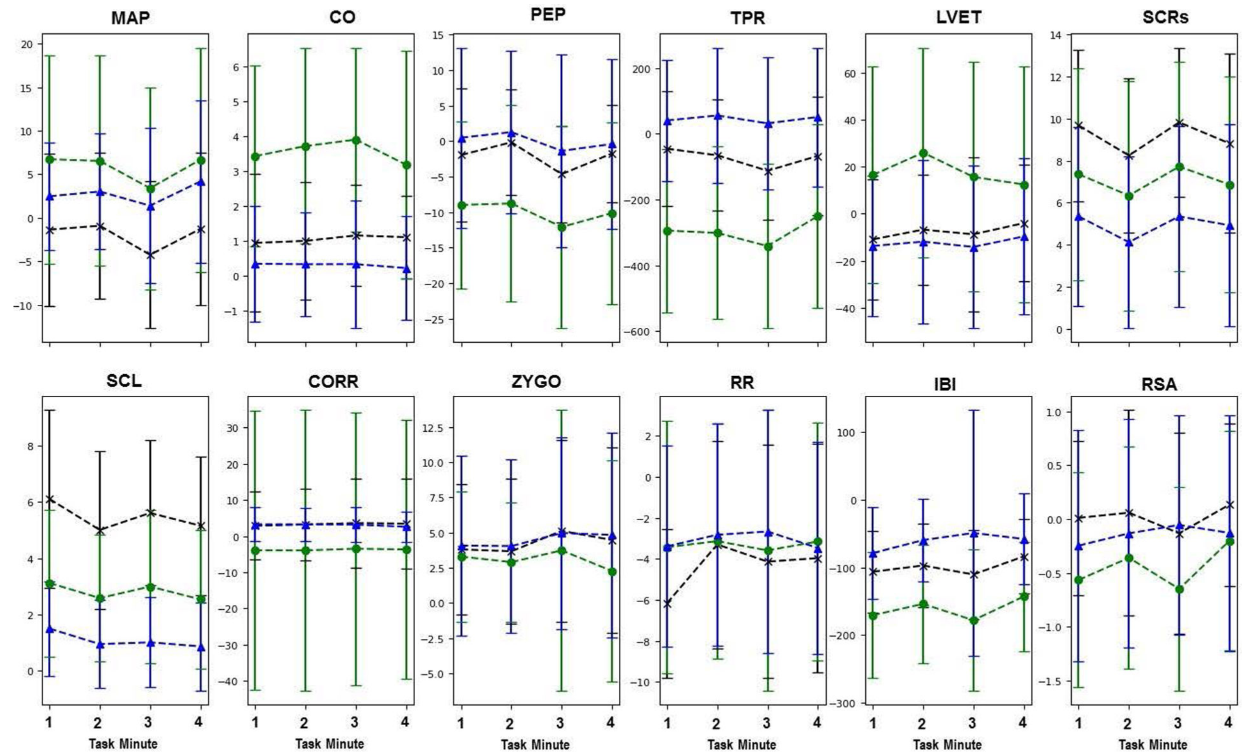

Figure 3.

Mean and standard deviation for each physiological feature for each minute of the task by cluster for Dataset 1 in the 3-cluster solution. The green cluster ($n = 61$) shows a challenge-like pattern of physiological activity, the black cluster ($n = 33$) shows a threat-like pattern of physiological activity, and the blue cluster ($n = 71$) shows a pattern of physiological non-response. SCRs represent the number of SCRs during each minute of the mental math task. All other variables are presented as change scores from baseline in feature-specific units: MAP is in mmHg; CO is in L/min; PEP, LVET, and IBI are in msec; TPR is in dyne·s·cm^{-5}, SCL is in microSiemens, CORR and ZYGO are in microvolts, and RR is in breaths/min.

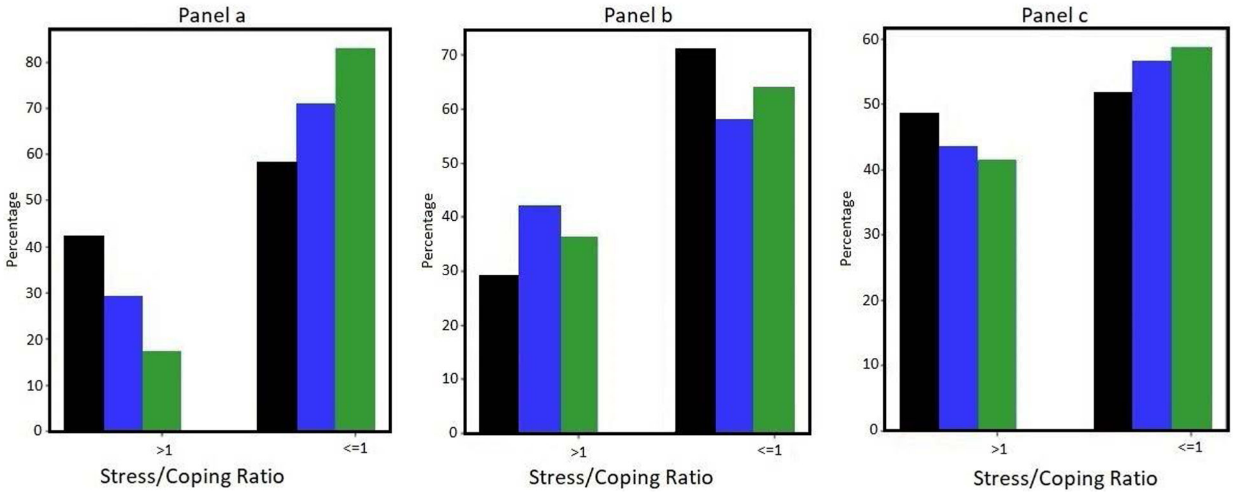

Figure 4.

Proportion of individuals in each of the three clusters with a stress/coping ratio indicating threat (>1) or challenge (< = 1) for Dataset 1 for self-reported appraisals provide before the task (Panel a), during the task (Panel b), and after the task (Panel c). Black bars represent individuals in the black “threat-like” cluster, green bars represent individuals in the green “challenge-like” cluster, and blue bars represent individuals in the blue “non-responder” cluster.

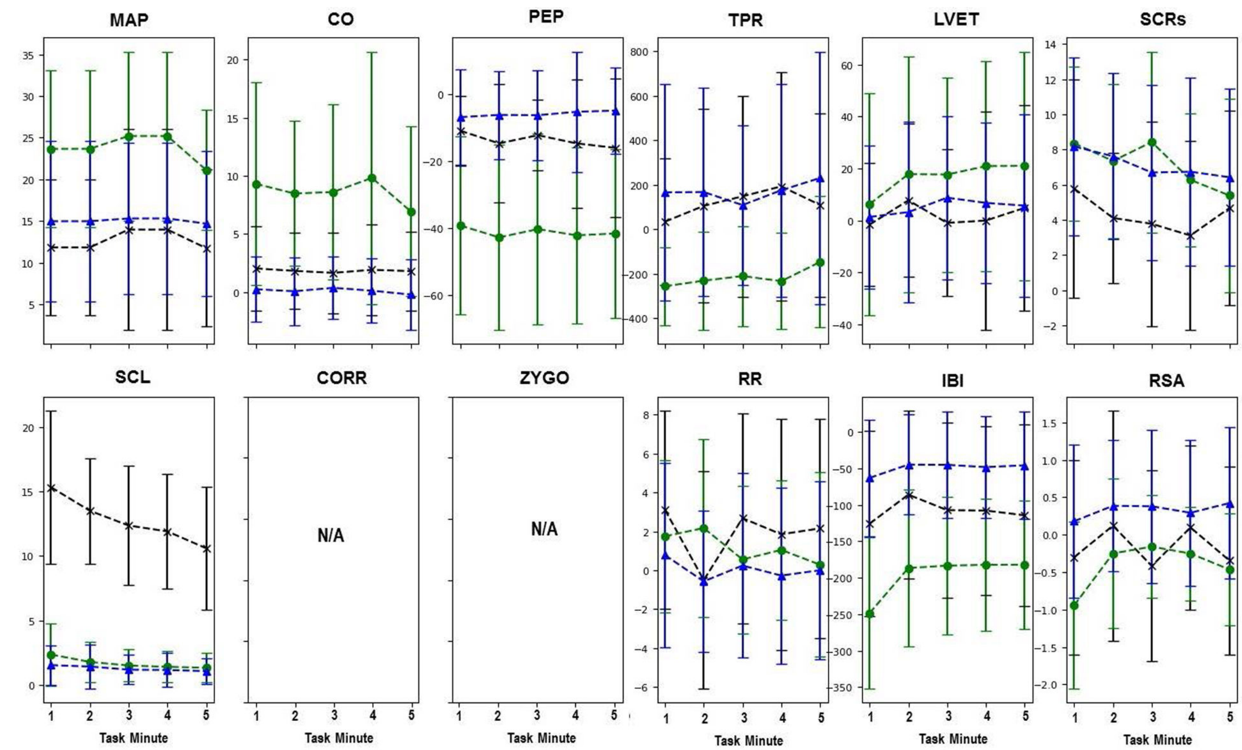

Figure 5.

Mean and standard deviation for each physiological feature for each minute of the task by cluster for Dataset 2 in the 3-cluster solution. The green cluster ($n = 13$) shows a challenge-like pattern of physiological activity, the black cluster ($n = 9$) shows a threat-like pattern of physiological activity, and the blue cluster ($n = 108$) shows a pattern of physiological non-response. SCRs represent the number of SCRs during each minute of the mental math task. All other variables are presented as change scores from baseline in feature-specific units: MAP is in mmHg; CO is in L/min; PEP, LVET, and IBI are in msec; TPR is in dyne·s·cm^{-5}, SCL is in microSiemens, and RR is in breaths/min.

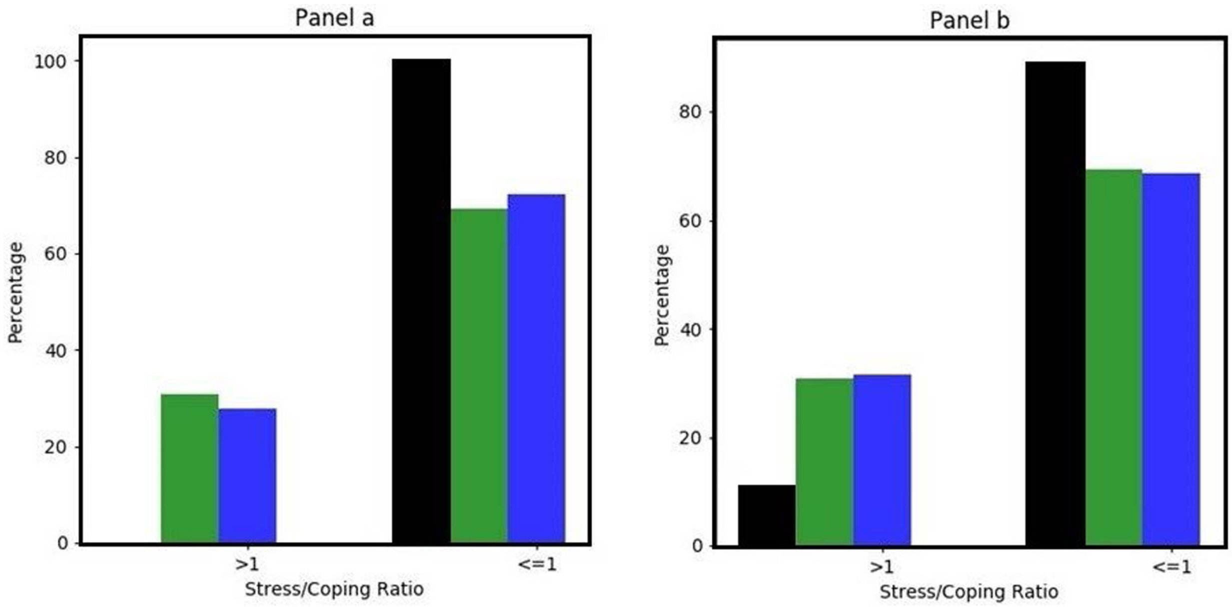

Figure 6.

Proportion of individuals in each of the three clusters with a stress/coping ratio indicating threat (>1) or challenge (< = 1) for Dataset 2 for self-reported appraisals provide before the task (panel a) and after the task (panel b). Black bars represent individuals in the black "threat-like" cluster, green bars represent individuals in the green "challenge-like" cluster, and blue bars represent individuals in the blue "non-responder" cluster.

Table 1.

Demographic Information and Baseline Physiology for Participants included (versus Excluded) in the Analyses by Dataset.

Measure	Dataset 1				Dataset 2			
	Excluded Sample (*N* = 95)		Included Sample (*N* = 165)		Excluded Sample (*N* = 169)		Included Sample (*N* = 130)	
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Gender	95	44% Male	164	35% Male	169	38.5% Male	130	38.5% Male
Race	**95**	**42.1% White***	**165**	**64.2% White***	**169**	**40.2% White***	**130**	**52.3% White***
Age	**95**	**26.87 (11.45)***	**164**	**23.57 (9.122)***	155	23.79 (8.34)	130	22.91 (7.64)
Height (cm)	95	170.76 (8.35)	164	169.80 (12.09)	134	169.80(9.53)	125	169.38 (8.67)
Weight (kg)	**95**	**73.31 (19.33)***	**164**	**67.86 (14.55)***	**133**	**72.76(15.33)***	**129**	**67.61 (15.50)***
LVET BL	64	258.92 (50.38)	165	264.11 (52.02)	66	273.77 (38.55)	130	274.87 (34.75)
CO BL	64	8.44 (3.97)	165	9.20(4.15)	66	7.50 (3.87)	130	9.16(6.52)
PEP BL	60	120.34 (20.76)	165	125.35 (17.70)	66	130.48 (18.49)	130	127.51 (16.44)
RR BL	86	14.32 (3.74)	165	14.67 (3.34)	97	16.28 (2.44)	130	16.42 (2.83)
IBI BL	86	943.07 (242.81)	165	913.41 (151.41)	97	848.69 (111.43)	130	855.19 (140.69)
RSA BL	85	6.51 (1.25)	165	6.68 (1.14)	97	6.28(1.18)	130	6.38(1.12)
SCR BL	92	2.99 (3.97)	165	2.36(3.60)	113	2.86 (3.69)	130	3.43 (3.42)
SCL BL	92	3.59 (4.70)	165	3.76 (4.06)	113	2.06 (4.60)	130	1.51(3.91)
MAP BL	81	83.15 (16.00)	165	84.82 (10.17)	124	81.47(8.51)	130	79.12 (8.07)
TPR BL	47	977.85 (493.95)	165	893.36 (453.54)	55	1067.48 (543.37)	130	909.77 (473.79)
CORR BL	91	6.52 (6.15)	165	7.83 (24.66)				
ZYGO BL	89	6.12 (10.26)	165	6.11 (10.03)				

Note: BL stands for baseline level. Independent samples t-tests were used to compare included and excluded participants from within each Dataset. Significant differences are printed in bold and marked with an asterisk (*p<.05). Equal variances were not assumed where appropriate. Compared to participants excluded from the present analyses, the included samples in both datasets contained a significantly larger proportion of White participants with a significantly lower average weight. In Dataset 1 only, the included sample was also significantly younger than the excluded sample, and in Dataset 2 only, the included sample had significantly lower mean arterial blood pressure and greater cardiac output at baseline than the excluded sample.

Table 2.

Goodness-of-Fit Measures (Silhouette Scores) for 2 to 6 Cluster Solutions for Both Datasets

Number of Clusters	Dataset 1	Dataset 2
2	.2832	.2632
3	.2134	.2592
4	.2127	.0944
5	.1785	.0901
6	.2006	.0590

Table 3.

Means, Standard Deviations, and Reliability Estimates for Physiological Reactivity Measures By Dataset.

Variable	Dataset 1			Dataset 2		
	Mean	SD	ICC	Mean	SD	ICC
LVET	−0.29	40.76	.72	6.14	33.63	.74
CO	1.65	2.62	.70	1.08	4.59	.83
PEP	−4.14	12.87	.75	−9.90	19.42	.80
RR	−3.44	5.66	.57	0.27	4.59	.50
IBI	−106.17	104.72	.67	−68.07	92.64	.91
RSA	−0.22	1.02	.72	0.22	1.04	.70
SCR	6.56	4.87	.75	6.94	5.13	.77
SCL	2.60	2.77	.90	2.10	3.48	.92
MAP	2.94	10.28	.86	15.75	9.69	.83
TPR	−105.39	268.18	.81	127.48	470.93	.87
CORR	0.59	23.91	.99	--	--	--
ZYGO	3.91	6.71	.62	--	--	--

Note: SD stands for standard deviation. ICC stands for intraclass correlation coefficient, a measure of consistency across trials within each measure.

Table 4.

Feature relevance scores as reflected in the scatter ratio or mutual information (shown in descending order) for the three-cluster solution for Datasets 1 and 2.

Dataset1				Dataset2			
Variables	**Scatter Ratio**	**Variables**	**Mutual Information**	**Variables**	**Scatter Ratio**	**Variables**	**Mutual Information**
SCL	.5251	TPR	.6178	SCL	2.3580	SCL	.5760
TPR	.4926	CO	.5964	CO	.4576	CO	.4964
CO	.4758	SCL	.5813	PEP	.4219	PEP	.4638
IBI	.3080	IBI	.4994	IBI	.3105	IBI	.4457
PEP	.1444	LVET	.4645	MAP	.0861	TPR	.3305
LVET	.1384	PEP	.3812	TPR	.0765	SCR	.1610
SCR	.1222	CORR	.3361	RSA	.0658	MAP	.1280
MAP	.0956	MAP	.2830	SCR	.0250	RSA	.1227
CORR	.0554	RMS	.1971	RR	.0219	RR	.0867
RSA	.0408	SCR	.1951	LVET	.0135	LVET	.0000
RR	.0142	RSA	.0725				
ZYGO	.0111	RR	.0262				

Table 5.

Demographic Information and Baseline Physiology for Participants by Cluster in the 3-Cluster Solution by Dataset.

Measure	Dataset 1			Dataset 2		
	Blue Cluster	**Green Cluster**	**Black Cluster**	**Blue Cluster**	**Green Cluster**	**Black Cluster**
N	71	61	33[+]	108	13	9
Gender	43.7% Male	29.5% Male	27.3% Male	36.1% Male	38.5% Male	66.7% Male
Race	**54.9% White***	**75.4% White***	**63.6% White**	50.9% White	46.2% White	77.8% White
Age	**25.94 (11.59)***	**21.85 (7.14)**	**21.59 (3.59)**	23.35 (8.20)	21.15(3.81)	20.33 (2.65)
Height (cm)	171.92 (10.96)	167.55 (13.42)	169.37 (11.30)	169.01 (8.37)	169.20 (7.26)	173.90(12.92)
Weight (kg)	69.48 (15.20)	66.16(12.39)	67.48 (16.78)	68.07(16.11)	59.68 (9.74)	73.58 (10.75)
LVET BL	**280.37 (48.34)***	**243.16 (49.74)**	**267.83 (51.78)**	275.73 (34.38)	269.98 (32.39)	271.56 (45.04)
CO BL	**9.99 (3.98)**	**7.94 (3.24)***	**9.84 (5.38)**	9.31(6.85)	8.22 (4.10)	8.75(4.39)
PEP BL	123.91 (17.66)	125.69 (16.83)	127.85 (19.51)	126.81 (16.64)	127.46 (12.36)	135.96 (18.44)
RR BL	**13.87 (3.61)***	**14.93 (3.12)**	**15.90 (2.69)***	16.33 (2.90)	17.02 (1.90)	16.60 (3.27)
IBI BL	901.46 (140.55)	917.94 (164.89)	930.73 (150.47)	843.29 (136.79)	910.15 (88.13)	918.55 (214.28)
RSA BL	**6.40** (1.18)*	**6.92 (1.11)**	**6.87 (0.98)**	6.30(1.07)	6.86 (0.94)	6.68(1.72)
SCR BL	2.49 (3.95)	2.64(3.75)	1.56(2.22)	3.16(3.43)	4.26 (2.72)	5.44 (3.67)
SCL BL	3.08 (4.19)	3.94(3.67)	4.88 (4.31)	**0.86** (1.32)*	**1.17(1.15)***	**9.69 (11.81)***
MAP BL	85.29 (8.13)	83.09 (13.04)	86.99 (7.45)	79.17 (8.15)	79.19 (8.40)	78.22 (7.47)
TPR BL	**790.80(385.25)***	**850.43 (329.16)**	**1035.96 (545.14)**	908.97 (488.93)	914.67 (316.45)	912.28 (451.97)
CORR BL	5.24(4.78)	6.86 (7.65)	5.90(3.71)			
ZYGO BL	5.13(6.33)	6.18(11.09)	8.10(13.81)			

Note: BL stands for baseline level. One-way ANOVAs were used to compare participants across clusters within each Dataset. Significant differences are printed in bold and marked with an asterisk (*p<. 05). Equal variances were not assumed where appropriate.

[+]One participant from the black cluster in Dataset 1 is missing height, weight, gender, and age data.

Table 6.

Performance and Appraisal Ratings by Cluster in the 3-Cluster Solution by Dataset.

	Dataset 1			Dataset 2		
	Blue Cluster	**Green Cluster**	**Black Cluster**	**Blue Cluster**	**Green Cluster**	**Black Cluster**
Stress Ratings						
Before Task	3.10(1.02)	3.10 (1.00)	3.00 (1.17)	5.39 (2.26)	5.31 (2.21)	3.67 (1.73)
During Task	3.30 (1.01)	3.09 (1.08)	3.27 (1.20)	–	–	–
After Task	3.59 (1.01)	3.25 (1.18)	3.40 (1.07)	5.89 (2.20)	5.15 (2.58)	3.78 (1.99)
Coping Ratings						
Before Task	3.87 (0.98)	3.81 (0.96)	3.73 (1.08)	6.74 (1.66)	7.38 (1.12)	7.67 (0.87)
During Task	3.75 (0.87)	3.67 (1.00)	3.63 (0.96)	–	–	–
After Task	3.48 (1.16)	3.63 (1.19)	3.27 (1.39)	6.45 (2.10)	6.62 (2.02)	7.89 (1.27)
Responses Attempted						
Mini	14.68 (7.28)	13.67 (7.04)	14.97 (7.38)	11.03 (6.47)	11.23 (5.33)	16.33 (5.32)
Min2	10.21 (5.43)	9.44 (4.08)	10.33 (6.19)	9.15 (5.31)	9.15 (4.30)	10.00 (3.97)
Min3[+]	11.96 (6.66)	12.41 (6.28)	12.09 (6.46)	11.25 (6.25)	11.23 (5.13)	13.11 (5.26)
Min 4	–	–	–	9.48 (5.70)	10.54 (4.08)	9.56 (6.11)
Min5	–	–	–	9.20(5.81)	11.15(4.41)	12.67 (7.02)
Proportion Correct						
Mini	0.80 (0.25)	0.80 (0.26)	0.83 (0.27)	0.69 (0.54)	0.65 (0.62)	0.97 (0.05)
Min 2	0.77 (0.32)	0.74 (0.35)	0.81 (0.33)	0.83 (0.24)	0.76 (0.31)	0.90 (0.08)
Min3[+]	0.81 (0.23)	0.78 (0.26)	0.76 (0.24)	0.85 (0.29)	0.88 (0.14)	0.90 (0.19)
Min 4	–	–	–	0.66 (0.60)	0.83 (0.25)	0.74 (0.37)
Min 5	–			0.70 (0.44)	0.71 (0.68)	0.81 (0.32)

Note: One-way ANOVAs were used to compare participants across clusters within each Dataset. Equal variances were not assumed where appropriate. No differences passed Bonferroni-corrected alpha levels for multiple comparisons. Proportion correct was calculated only from participants who attempted at least 1 response.

[+] For Dataset 1, responses attempted and proportion correct in Minute 3 reflect performance in the third trial of the task (minutes 3 and 4 combined).